Comparison of land skin temperature from a land model, remote sensing, and in-situ measurement

Aihui Wang[1*], Michael Barlage[2], Xubin Zeng[3], Clara Sophie Draper[4]

[1] Nansen-Zhu International Research Centre, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China.
[2] Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA
[3] Department of Atmospheric Sciences, The University of Arizona, Tucson, Arizona, USA
[4] Global Modeling and Assimilation Office, NASA GSFC, Greenbelt, Maryland, USA and GESTAR / Universities Space Research Association, Columbia, Maryland, USA

Journal of Geophysical Research-Atmospheres
(submitted on 10/10/2013)

* Corresponding author address:
Aihui Wang
Nansen-Zhu International Research Centre
Institute of Atmospheric Physics
Chinese Academy of Sciences
P.O. Box 9804, Beijing 100029
P. R. China
wangaihui@mail.iap.ac.cn

Abstract

Land skin temperature (Ts) is an important parameter in the energy exchange between the land surface and atmosphere. Here hourly Ts from the Community Land Model Version 4.0, MODIS satellite observations, and in-situ observations in 2003 were compared. Compared with the in-situ observations over four semi-arid stations, both MODIS and modeled Ts show negative biases, but MODIS shows an overall better performance. Global distribution of differences between MODIS and modeled Ts shows diurnal, seasonal, and spatial variations. Over sparsely vegetated areas, the model Ts is generally lower than the MODIS observed Ts during the daytime, while the situation is opposite at nighttime. The revision of roughness length for heat and the constraint of minimum friction velocity from Zeng et al. [2012] bring the modeled Ts closer to MODIS during the day, and have little effect on Ts at night. Five factors contributing to the Ts differences between the model and MODIS are identified, including the difficulty in properly accounting for cloud cover information at the appropriate temporal and spatial resolutions, and uncertainties in surface energy balance computation, atmospheric forcing data, surface emissivity, and MODIS Ts data. These findings have implications for the cross-evaluation of modeled and remotely sensed Ts, as well as the data assimilation of Ts observations into Earth system models.

1. Introduction

Land skin temperature (Ts) is one of the key variables of the earth system, acting as the lower boundary of the atmosphere. The difference between Ts and overlying atmospheric temperature (Ta) helps to determine the partitioning of surface energy fluxes into sensible and latent heat fluxes [*Garratt*, 1995; *Prigent et al*., 2003]. Ts also controls the amount of heat transfer from the land surface into the soil, and then indirectly affects thermal states in deep soil. Hence, there is potential to improve land surface flux forecasts by assimilating Ts observations [e.g., *Bosilovich et al.*, 2007; *Ghent et al.*, 2010; *Reichle et al.*, 2010; *Xu et al*., 2011]. Although the importance of Ts has been recognized, the accuracy of global Ts datasets over land is not well understood.

Land surface models (LSMs) driven by observation-based atmospheric data are widely used to produce Ts. The upward longwave radiation fluxes simulated by LSMs combined with downward longwave radiation fluxes and the surface emissivity can be used to estimate long-term high resolution Ts continuously. Solar radiation is a driving force of Ts, which is evident in clearly correlated diurnal and seasonal variations. The magnitude of modeled Ts is affected by surface land cover, soil moisture, and soil properties (e.g., soil albedo and soil texture). Due to large land surface heterogeneities, energy fluxes are difficult to simulate accurately in LSMs. Even over a bare ground grid cell, LSMs still have difficulty in realistically producing skin temperature and surface fluxes [*Chen et al*., 2010; *Zheng et al*., 2012; *Zeng et al.*, 2012]. Efforts have also been made to improve the simulation of Ts in LSMs. For example, the

underestimation of diurnal Ts variation over the Tibetan Plateau is a notable deficiency in most LSMs due to the incorrectly parameterized roughness length for heat (z_{oh}). *Yang et al.* [2002] developed a new z_{oh} formulation from observations at the Tibetan Plateau to improve surface turbulence flux parameterization over bare soil surface, which also improved the Ts simulation in the Noah LSM [*Chen et al.*, 2010]. Based on theoretical arguments and synthesis of previous observational and modeling efforts, *Zeng et al.* [2012] improved the Ts diurnal range simulated over bare ground in two LSMs through z_{oh} revisions, constraining minimum friction velocity, and modification of soil thermal conductivity. *Zheng et al.* [2012] adopted a new vegetation-dependent formulation of momentum and thermal roughness lengths in the National Center for Environmental Prediction (NCEP) Global Forecast System (GFS), and substantially reduced the cold forecast bias during the day, which then improved the brightness temperature in the NCEP data assimilation system.

Many previous evaluation and validation studies involving Ts modeling have been based on single point station measurements. However, Ts is not a routinely measured variable at meteorological stations, and it is only available at a very limited number of stations with relatively short data records [e.g., *Augustine et al.*, 2000; *Baldocchi et al.,* 2001]. Satellite observations can produce land surface measurements over large areas with high spatial resolutions. For example, global clear-sky Ts products from the Moderate Resolution Imaging Spectra-radiometer [MODIS, *Salomonson et al.*, 1989] have been available since 2000. The MODIS sensor provides a quality data source of Ts for model evaluation from four daily satellite

overpasses [e.g., *Ghent et al.*, 2010] and for data assimilation [e.g., *Bosilovich et al.*, 2007; *Reichle et al.*, 2010; *Xu et al*., 2011].

In this study, through comparisons of Ts from the Community Land Model version 4.0 (CLM4) with both the MODIS (globally) and in-situ station measurements (at four locations), we test whether the differences between monthly mean Ts from these three data sources can be used to better identify errors in, and hence make improvements to, either of the modeled or remotely sensed data sets. At the same time, in order to improve the global Ts simulation over bare soil surfaces, the new parameterization schemes in *Zeng et al*. [2012] were implemented into CLM4.0. Comparing these three data sets is not straightforward, since substantial representative differences are expected between Ts estimates obtained from in situ sensors, remote sensors, and land surface models, most notably due to the differences in the typical spatial resolution of each of these estimates.

Section 2 introduces the MODIS Ts, while Section 3 describes the computations of Ts in CLM4.0 and the modification of parameterizations. Results are presented in Section 4, and a summary is given in Section 5.

2. MODIS skin temperature

Two MODIS instruments were installed on the NASA Terra and Aqua satellite platforms, which were launched in December 1999 and May 2002, respectively. Aqua overpasses around local solar time of 1:30pm (ascending mode) and 1:30am

(descending mode), while Terra is around 10:30am (descending mode) and 10:30pm (ascending mode). The global 0.05°x0.05° spatial resolution monthly MODIS collection 5 Ts data (MODIS product name: MOD11C3/MYD11C3) used in this work were retrieved from the thermal infrared (TIR) bands using the generalized spilt-window algorithm [*Wan et al*., 2008]. Since the surface TIR signal is difficult to determine with the presence of clouds, the MODIS monthly Ts product includes information on the individual cloud covered days that were used to filter out cloud-contaminated observations when calculating the mean monthly observed (in situ or remotely sensed) Ts.

The accuracy of satellite Ts is affected by surface retrieval techniques, cloud condition, and land surface properties [*Wan et al.*, 2004; 2008], which all significantly constrain the application range of such products. Therefore, evaluation and validation of remote sensing products based on ground-measurement values are important and necessary [e.g., *Wan et al.*, 2002; 2004; 2008; *Wang and Liang*, 2009; *Zheng et al.*, 2012]. For example, *Wan et al*. [2004] used the observed data over 20 stations to validate the MODIS Ts. *Wang and Liang* [2009] evaluated the MODIS Ts with six Surface Radiation Budget Monitoring stations [SURFARD, *Augustine et al.*, 2000]. Studies such as these are essential to understanding the application capability and accuracy of satellite observed Ts.

3. Skin temperature in CLM4.0

CLM4.0 is the land component of the Community Earth System Model (CESM), and it can also be used as a stand-alone model to simulate the land surface heat and hydrological variables [*Lawrence et al.*, 2012], as used here. Compared with earlier versions of the model, CLM4.0 has several important modifications and has implemented additional components, including updates to soil hydrology, soil thermodynamics, albedo parameters, a carbon–nitrogen biogeochemical model, an urban canyon model, as well as revised soil and snow sub-models [*Oleson et al.*, 2010; *Lawrence et al.*, 2011]. The surface skin temperature Ts for a model grid box is not explicitly computed in CLM4.0, but it can be derived from the surface incoming (LW↓) and outgoing (LW↑) longwave radiation combined with surface emissivity (ε):

$$\varepsilon\sigma T_s^4 = \mathrm{LW}\uparrow - \mathrm{LW}\downarrow \cdot (1-\varepsilon), \qquad (1)$$

where $\sigma=5.67\times10^{8}\mathrm{Wm}^{-2}\mathrm{K}^{-4}$ is the Stefan-Boltzmann constant. In CLM4.0, surface emissivity over non-vegetated surfaces is constant: 0.96 for soil and wetland, and 0.97 for glacier. Over vegetated surfaces, surface emissivity (ε_v) is a function of the leaf (L) and stem area index (S):

$$\varepsilon_v = 1 - e^{-(L+S)/\bar{\mu}}, \qquad (2)$$

where $\bar{\mu} = 1$ is the average inverse optical depth for longwave radiation. The grid box in CLM4.0 is a hybrid of different land unit types (e.g., bare soil, vegetation, glacier, wetland, and urban). Over the vegetated part of a grid cell, the vegetation can be described by up to 16 unique vegetation categories [*Oleson et al.*, 2010]. The grid box averaged LW↑ in the model is computed from the areal weighted LW↑ from both vegetated and bare ground areas.

It has been widely recognized that z_{oh} is important in the parameterization of surface fluxes [Zeng and Dickinson, 1998; *Yang et al.*, 2002, 2008; *Zeng et al.*, 2012]. In LSMs, z_{oh} is usually a function of roughness length of momentum (z_{om}) for bare surfaces, or proportional to the canopy height for the vegetated surfaces [*Zeng and Dickinson*, 1998; *Oleson et al*., 2010]. However, using the current z_{oh} scheme, CLM substantially underestimates diurnal variations of Ts, similar to other LSMs [*Chen et al*., 2010; *Zeng et al*, 2012; *Zheng et al.*, 2012]. Through both theoretical analyses and data-model comparison, *Zeng et al*. [2012] suggested some revisions for the model parameterization schemes that substantially improved Ts simulations over two semi-arid sites in both CLM3.5 and the Noah LSMs. Here, we extend those modifications to global CLM4.0 simulations, and we simply describe the new parameterization schemes.

Zeng et al. [2012] modified the z_{oh} formulation

$$\ln\left(\frac{z_{om}}{z_{oh}}\right) = a\left(\frac{u_* z_{om}}{\nu}\right)^{b} \qquad (3)$$

where $\nu = 1.5\times10^{-5}$ m^2s^{-1} is the molecular viscosity, $b = 0.5$ and $a = 0.36$. These values are 0.45 and 0.13 in the default CLM4.0, respectively.

Another model deficiency is that under stable conditions (usually during nighttime) the computed sensible heat is near zero and largely underestimated, which leads to the decoupling of atmospheric boundary layer from the land surface [*Beljaars and Viterbo*, 1998]. *Zeng et al.* [2012] also suggested constraining the minimum friction velocity under stable condition,

$$u_{*min} = 0.07\frac{\rho_o}{\rho}\left(\frac{z_{om}}{z_{og}}\right)^{0.18} \qquad (4)$$

where ρ (ρ_0) is the air density at reference (sea) level, z_{om} is surface roughness length for momentum; and z_{og} = 0.01 m is the roughness length of bare soil. A similar method has been widely used in eddy-correlation flux measurements from towers [*Gu et al*., 2005]. Because air density correlates with the terrain height, equation (4) implicitly considers the elevation effects in the computation of sensible heat. Equation (4) is not used in the default CLM4.0.

In the modeling experiments presented below, CLM4.0 was run offline at a 1.9°x2.5° horizontal resolution driven by an observation-based global atmospheric forcing dataset [*Qian et al.*, 2006]. Other parameters, such as vegetation parameters and soil properties, are from the standard model data package [*Oleson et al*., 2010]. The model was run for 1995-2004, with the multi-year "spun-up" initialization [*Lawrence et al*., 2012], and the results in 2003 were analyzed and compared with both observations and satellites products.

Two model experiments were conducted: one with the default model parameterization referred to as CLM-C, and another with modifications described by equations (3) and (4) denoted as CLM-N. The hourly outputs of LW↑and surface emissivity combined with LW↓ in the atmospheric forcing dataset were used to compute Ts from equation (1) over global land areas. In order to compare with MODIS, the modeled Ts was interpolated to the four MODIS satellite overpass times.

4. Results

4.1 Comparisons of Ts from CLM4.0, MODIS, and in-situ measurements

Ground measurements at four stations with barren-dominant land cover types are used to compare with both MODIS and CLM4.0 simulations. Based on equation (1), Ts at each station was computed from the measurements of surface incoming and emitted LW, combined with surface emissivity (Table 1). Using inverse-distance-weighted interpolation method, MODIS Ts over each station was interpolated from four 0.05°closest pixels, and the modeled Ts was also interpolated from four closest model grid boxes. The monthly mean Ts was computed using only the days that were observed as clear sky by MODIS. For example, over Desert Rock at 1:30pm, there were 25 days in July 2003 under clear sky conditions, and the monthly mean in situ and modeled Ts values were calculated using only those 25 days.

Table 2 compares the monthly mean Ts differences at four times over four stations between MODIS and CLM4.0 simulations versus in-situ observations in July 2003. These stations over dry regions show large diurnal variations. For example, the monthly averaged Ts differences between 1:30pm and 1:30am under clear sky conditions from in-situ measurements are 29.9 K, 27.2 K, 17.22 K, and 25.18 K over Desert Rock, Colorado, Tongyu, and Gaize, respectively. Both MODIS and modeled Ts show negative mean differences (MDs) compared with the in situ data (i.e., are cooler than in situ Ts) at most times at all four stations, and most MDs are statistically significant at 1% level (Table 2). Both CLM-C and CLM-N have large negative MDs

(up to -11.41K for CLM-C and -8.91K for CLM-N, both at 1:30pm at Gaize). MODIS has negative MDs at night, ranging from -1.93K (10:30pm at Tongyu) to -5.21K (10:30pm at Gaize), while its MDs could be positive or negative during daytime, ranging from -2.30K (10:30am at Tongyu) to 10.61K (10:30am at Gaize). If the abnormally high MD at 10:30am at Gaize is excluded, the daytime MODIS MDs are generally smaller in magnitude than nighttime values.

The root mean square difference between the different Ts data sets used here would be dominated by these large MDs. However, these MD are not necessarily due to errors in a specific data set, and may be due to representative differences between them (e.g., differences in the spatial resolution, including potentially the land cover, between the data sets). Therefore, we compute the standard derivation of differences (STDd) between model or MODIS results and in situ observations. The STDd is the root mean square difference between the data sets, once the bias between them has been removed. Recognizing the different standard deviations of the in situ data (STDo) between daytime and nighttime, Table 3 shows that the ratios of STDd/STDo vary from 0.50 to 1.81 for MODIS, 0.20-1.18 for CLM-C, and 0.20-1.33 for CLM-N. These ratios are on average greatest at 10:30pm for MODIS and at 1:30pm for CLM-C and CLM-N.

Among the 16 MD values in each column of Table 2, 11 (or 10) values from MODIS are smaller in magnitude than those from CLM-C (or CLM-N). On the other hand, 11 of the 16 ratios in Table 3 from CLM-C and CLM-N are smaller than those from MODIS. CLM-N has 15 values smaller in magnitude than CLM-C in Table 2,

demonstrating the improvement in CLM-N, while 13 of the 16 ratios in Table 3 are within 0.02 between CLM-C and CLM-N.

While the better performance of MODIS data than model results in terms of MDs to the in-situ data is expected, it is still surprising to see the much larger MODIS MDs (in magnitude) in Table 2 than reported in previous studies [*Wan et al.*, 2002, 2004, 2008]. For example, *Wan et al*. [2004] indicated that the Ts biases of MODIS from station observations are within 1 K. A potential reason is that previous validation studies used the MODIS Ts data at the highest resolution (1 km) under clear-sky conditions while we use the MODIS data at 0.05° (~5 km) grid cells for global studies. In general, our 5 km MODIS Ts data used in Tables 2 and 3 may contain partially cloudy conditions and hence contain more days of data in a given month. For instance, at 10:30pm at Gaize, while the MODIS MD is -5.21 K in July, it is less than 1.45 K in magnitude for 25% of the days. On the other hand, at 10:30am at Gaize, the MODIS MD is 10.61 K in July, and such a large positive bias indicates the deficiency of the MODIS data at this time over this high altitude location.

It is also interesting to note that MODIS from Aqua (1:30am/1:30pm) performs better than that from Terra (10:30am/10:30pm) compared with in situ measurements. For nighttime MDs (at 1:30am and 10:30pm) and daytime values (at 1:30pm and 10:30am) in Table 2, Aqua (or Terra) gives smaller MDs in magnitude seven (or just one) times. Similarly, Aqua (or Terra) gives smaller ratios seven (or just one) times in Table 3.

The MDs of CLM-N in Table 2 are also much larger than those reported in *Zeng*

et al. [2012] at both Desert Rock and Gaize sites. The improvement of daytime Ts in CLM-N over CLM-C is substantial in *Zeng et al.* [2012], while it is more moderate in Table 2. These different results can be reconciled along several different lines. In the results presented here, the model was run globally at coarse resolution (1.9°x2.5°) where only 65% of the grid box near Desert Rock was of the bare soil, while in *Zeng et al*. [2012] CLM4.0 was run at a single point with 100% bare soil fraction at this site. Furthermore, the atmospheric forcing data, particularly air temperature (which is related to elevation) and downward solar radiation (SWd), are very different between our simulations based on the *Qian et al.* [2006] data, and the in situ measurements used in *Zeng et al*. [2012]. For instance, Table 2 shows that 12 of the 16 air temperature differences between *Qian et al.* [2006] and in situ data are less than -3 K, and all SWd differences are negative. While some of these differences in the atmospheric forcing are due to errors in each data set, the large difference in spatial resolution of each atmospheric data set would also introduce some differences.

4.2 Evaluation of the CLM4.0 modeling with MODIS Ts

Using the 0.05° MODIS Ts data to evaluate global model output is not straightforward, and involves several steps. First, at each satellite overpass time (four times daily), MODIS monthly Ts data are spatially averaged within each CLM4.0 grid box with the requirement that at least 20% of the model grid box is defined as land in MODIS. Each 1.9°x2.5° CLM4.0 grid box potentially includes 1900 0.05°x0.05°

MODIS observations. Another important consideration is the potential for cloud contamination adversely affecting MODIS Ts. Scarino et al. (2013) found increased agreement between remotely sensed and in-situ Ts with decreasing cloud cover. The number of MODIS grid cells observed as clear-sky in each model grid box varies with month and location. Hence, we also calculate the clear sky fraction (CF) as the percentage of MODIS grid cells within each CLM4.0 grid box that are declared as clear on a given day. The CF values for an individual day averaged over global land (excluding the Antarctic) vary from 45-60%, and the monthly mean values in July are a little bit larger than in January.

Figure 1 shows the distribution of global CLM4.0 grid boxes based on the monthly mean of MODIS daily CF, binned into 10% intervals from 0-100%, at each satellite overpass time in January and July 2003, respectively. The clear-sky fraction is greater than 90% for ~25% of the model grid boxes in January and ~28% in July, primarily over semi-arid and arid regions, e.g., northern Africa, Middle East, western China, western and central Australia, and southwestern United States. CF is less than 10% for ~25% of model grid boxes in January and ~20% in July, primarily over tropical rainforests such as the Amazonia, equatorial Africa, and south-eastern Asia. The higher percentages of model grid boxes at the low CF bin in January (Figure 1a) than in July (Figure 1b) are related to the more extensive cloud cover in the wet season (including January) over tropical rainforests. For the four overpass times, the percentage of model grid boxes with CF $< 10\%$ is highest at 1:30am, consistent with the nighttime precipitation maximum over rainforests [*Angelis et al*., 2004]. The

percentage of model grid boxes for CF > 90% in July (Figure 1b) is higher during the day (at 10:30am and 1:30pm) than at night (at 1:30am and 10:30pm), probably because of the higher relative humidity at night over dry regions. For CF between 10-90%, the percentage of model grid boxes varies from 8.7-5.1%, and in the same CF bin they change little with satellite overpass times.

Since MODIS Ts observations are for clear-sky conditions only, the model Ts must also be screened for cloudy conditions before being compared to MODIS observed values. This screening is complicated by the spatial and temporal aggregation between the observed Ts and monthly mean modeled values. To address this we first bin the daily MODIS CF values for all model grid cells into 10% intervals from 0-100%. We then calculate the monthly model Ts for each bin from hourly CLM4.0 Ts from every day of the month. That is, for different daily CF bins, the number of grid boxes used to compute the monthly mean is different. For example, in July over northern hemisphere (NH), about 75% of the model land grid boxes were used in the computation of monthly mean values at 1:30pm for CF > 50%, while only about 50% of the model land grid boxes were used for CF > 90%. Furthermore, we require that the daily MODIS Ts data at each overpass time are available for at least 10 (clear) days in a month for the calculation of the monthly mean. Using these criteria, it is found that the MD between the modeled and remotely sensed Ts (i.e., mean CLM minus MODIS Ts) generally decrease with increasing CF values over both hemispheres. For instance, at 1:30 pm in July 2003, the MD over Southern Hemisphere (SH) land areas varies from 0.59 K (for CF<10%) to -0.32 K (for

CF>90%), while over NH, it varies from 1.55 K (for CF<10%) to 1.04 K (for CF>90%).

For CF > 50%, Figure 2 shows the spatial distribution of the differences between CLM-C and MODIS Ts at four satellite overpass times averaged in July 2003, respectively. The Ts biases display large spatial and diurnal variations, and the magnitudes of differences are substantial over some regions. At daytime, areas with negative biases are mainly located at midlatitude arid and semi-arid regions, while at nighttime positive biases are dominant over most of land areas. The global mean difference in NH varies from -2.17 K (at10:30am) to 4.33 K (10:30 pm), while in SH it varies between -2.40 K (10:30am) to 4.09 K (10:30pm). At 1:30pm, the mean differences over two hemispheres are smallest in magnitude among all four times, with values of 0.07 K in SH and 1.25 K in NH, respectively. *Wan et al.* [2004] also found that MODIS Ts at 1:30 pm is closer to in-situ measurements, and suggested that Ts at 1:30pm would be more suitable for climate change studies since the 1:30pm local solar time is closer to the maximum temperature of the land surface.

The mean differences between CLM-C and MODIS in January are on average larger in magnitude than those in July. For instance, the mean difference is 5.80 K in SH (versus 4.09 K in July in Figure 2). At 1:30pm, the mean difference over SH of 0.18 K is also the smallest in magnitude among all four times over both hemispheres. In January, due to the snow existence over northern high latitudes (and some midlatitude regions), satellite-retrieved surface products might contain large errors, and the comparison of CLM4.0 and MODIS data may not be appropriate.

4.3 Performance of the CLM4.0 with Eqs. (3) and (4)

As mentioned in *Zeng et al.* [2012], equation (3) primarily increases the daytime Ts with a negligible effect on nighttime Ts. Equation (4) slightly increases Ts under very weak wind and stable conditions at night. Figure 3 shows the Ts differences between CLM-N and CLM-C with respect to bare ground fractions in 5% intervals. Indeed the Ts from CLM-N is overall larger than that from CLM-C, and their differences increase with the bare soil fraction. The difference is more pronounced in the day time than at night and in the warm month than in the cold month. The largest difference is at 1:30 pm in January over SH, and the values are up to 6 K over the totally bare covered regions.

Therefore, we mainly focus on the evaluation of equations (3) and (4) with MODIS at day time over regions where bare ground fraction is greater than 30%. These regions include most of semi-arid and arid areas, such as northern Africa, Middle East, northwest China, Tibetan Plateau, central and western Australia, and small areas of southwestern United States.

Figures 4 and 5 plot the global distribution of Ts differences between CLM-N and CLM-C, and between CLM-C and MODIS at day times. The Ts differences vary seasonally and spatially, and they are greater in July than in January. The Ts differences between CLM-C and MODIS are generally negative over most regions, and they are less than -8 K (i.e., greater than 8K in magnitude) at 10:30am over part

of the northern China, Arabian Peninsula, and Sahara Desert (Figures 4c and 5c). The differences between CLM-N and CLM-C are positive over most regions, and CLM-N overall reduces the cold biases compared to MODIS Ts of CLM-C at day time shown in Figure 4.

Table 4 summarizes the hemisphere averaged results from Figures 4 and 5. The differences are all negative except at 1:30pm in July over NH between CLM-N and MODIS. This issue will be further discussed in section 4.4. The mean differences between CLM-N and MODIS are generally smaller than those between CLM-C and MODIS, suggesting that equations (3) and (4) reduce the cold bias of CLM-C.

4.4. Possible reasons for Ts biases between CLM4.0 and MODIS

The large differences between the Ts estimates from CLM4.0 and MODIS could be due to errors in either data set, or representative differences between them. With no independent measure of Ts at global scales, it is difficult to definitively attribute a cause to the large mean differences obtained above. However, cross-referencing these mean differences with independent information on the accuracy of each data set can help to confirm known problems in each data source.

For the large Ts differences in Figure 4 and Table 4, we can identify several possible reasons. First, there are deficiencies in the energy balance computation in CLM4.0. In the past few years, many efforts have been made to reduce such deficiencies [*Zeng and Wang*, 2007; *Wang and Zeng*, 2009; *Zeng et al.*, 2012].

Equations (3) and (4) from *Zeng et al.* [2012] are also among such efforts. Indeed Table 4 shows that these revisions reduce the cold bias of CLM-C (compared to MODIS).

Second, there are deficiencies in the atmospheric forcing data [e.g., *Guo et al.*, 2006; *Wang and Zeng*, 2011]. For global land areas, accurate atmospheric forcing data are not available. The current global forcing data sets are usually based on reanalysis datasets with bias correction by limited in-situ or remote-sensed observations [e.g., *Qian et al.*, 2006; *Sheffield et al.*, 2006]. *Wang and Zeng* [2011] found that the precipitation and air temperature in the atmosphere forcing data of *Qian et al*. [2006] used in CLM4.0 are largely biased compared with in situ observation-based data over China, and these biases affect the modeled soil hydrology variables. As mentioned earlier, there are also large biases, compare to in situ data, in the air temperature and downward solar radiation in the forcing data of *Qian et al.* [2006] (Table 2), which are likely in part due to differences in spatial resolution and elevation.

Furthermore, the Ts differences between CLM4.0 and MODIS are partially affected by the different treatment of surface emissivity in the Ts computation in equation (1). Surface emissivity is constant over bare soil and is a simple function of vegetation leaf area index in CLM4.0 (equation 2), while the MODIS surface emissivity is estimated from land cover type in each 0.05° pixel through MODIS thermal infrared (TIR) bands and a classification –based emissivity model [*Snyder et al.* 1998]. *Wan et al.* [2004] pointed out that errors in the classification –based

emissivity may be larger over semi-arid and arid regions due to larger temporal and spatial variations. Surface emissivity over bare soil is affected by many factors (e.g., surface chemical composition) and the wavelength at which the emissivity is measured [*Van De Griend and Owe*, 1993; *Jin and Liang*, 2006]. In particular, *Jin and Liang* [2006] found that assuming a constant surface emissivity over bare soil would strongly affect Ts and sensible heat fluxes over desert.

As mentioned earlier, CLM4.0 results represent the effective Ts over all land cover types present in each 1.9°x2.5° grid box, while the MODIS monthly Ts is computed from only the clear-sky 0.05° pixels in each grid box. We only used the days with MODIS clear sky fraction greater than 50% in each model grid box when we computed the monthly average of the modeled Ts in Figures 3-5. This means that we essentially compared the clear-sky MODIS Ts with model Ts under partially cloudy conditions. Since clouds decrease downward solar radiation, this would introduce a cold bias of daytime Ts between CLM4.0 and MODIS. On the other hand, if we only consider days with MODIS clear-sky fraction > 90% in each 1.9°x2.5° box, then the percentage of grid boxes would be about 30% in July and less than 25% in January (Figure 1), and the number of such days in each grid box would be also very limited.

Besides the model-observation inconsistencies and model forcing data deficiencies, there are also MODIS Ts deficiencies. As shown in Figure 4d and Table 4 at 1:30pm in July, the substantial positive biases between CLM-C and MODIS in northeastern Africa are opposite to those over other regions of the Sahara Desert. To

further explore this issue, we selected two grid boxes (centered around 29°N/23°E and 29°N/10°W) in northeastern Africa. Figure 6 shows that both CLM-C and MODIS have strong diurnal variations at both grid boxes. Because both boxes are located at the same latitude over the Sahara Desert with bare soil fraction greater than 90%, the differences of Ts (including its diurnal cycle) between the two boxes are expected to be smooth with time. This is indeed the case for CLM-C (Figure 6b). However, the MODIS Ts differences between two grid boxes show much more pronounced diurnal variations which also differ from month to month. Therefore, the abnormal positive Ts differences between CLM4.0 and MODIS over northeastern Africa are thought to be caused by MODIS Ts warm biases, which in turn may be related to MODIS surface emissivity deficiencies over this area [e.g., *Wan et al*., 2004]. A similar situation to Figure 6 was also found in the Amazon rainforest (Figure not shown), and the MODIS Ts bias there may be related to the difficulty in identifying clear-sky pixels.

5. Summary and further discussions

Land skin temperature (Ts) is one of the important parameters in the energy exchange between the land surface and atmosphere. Lack of global long-term in-situ Ts observations is a barrier to understanding the earth system. Land surface models and satellites provide two alternative ways to produce Ts. Various data sources, however, contain deficiencies and limitations, and their comparison would provide

some insights for the data developers and users.

In this study, Ts from MODIS, in-situ station measurements, and the Community Land Model version 4 (CLM4.0) simulations in 2003 were compared. Two modifications (i.e., equations (3) and (4)) are also implemented into CLM4.0. Hourly outputs of surface emitted longwave radiation combined with the surface downward thermal radiation fluxes are used to compute Ts over global land areas. MODIS Ts is only available during cloud-free conditions, while modeled Ts is the averaged-value of whole grid box regardless of cloud cover. Therefore, in the comparison of modeled and MODIS Ts, the MODIS clear-sky information is used to make the comparison more consistent.

Results show that both MODIS and modeled Ts datasets can capture the diurnal variation of Ts at four station locations, but also display distinct biases compared to the in situ data. Both MODIS and modeled Ts show significant negative mean differences at most times in July 2003, and the mean differences are statistically significant at the 1% level. The magnitude of biases varies by station and time. The MODIS Ts is generally closer to station observations than the model simulations are.

Under the 50% MODIS clear-sky fraction conditions, global comparisons between the MODIS and modeled Ts also show that their mean differences vary spatially and seasonally. Over land areas the mean differences are mostly negative during the day (i.e., model has a cold bias compared to MODIS) and positive at night. The modified CLM4.0 reduces this cold bias in the daytime over bare ground dominated regions.

Comparison of (TIR) remotely sensed and modeled Ts requires the consistent treatment of cloudy conditions between the two data sets, including in the calculation of spatially and/or temporally aggregated values. Furthermore, comparison of MODIS and modeled Ts can help to identify deficiencies in MODIS Ts over some regions, such as the Sahara Desert.

While the monthly mean time scale of this study is not directly relevant to most data assimilation applications, this work has some obvious implications for the assimilation of remotely sensed Ts into Earth system models. Most notably, the large biases between modeled and remotely sensed Ts are not unique to this study [e.g., *Ghent et al.*, 2010; *Scarino et al.*, 2013], and must be addressed before Ts data can be assimilated (since standard data assimilation techniques are contingent on the observations and the model being bias free). This is usually achieved by rescaling the observations to be consistent with the model Ts prior to assimilation [e.g., *Ghent et al.*, 2010; *Reichle et al.*, 2010]. Additionally, the need to carefully account for cloudy conditions when comparing modeled and observed Ts also applies to the assimilation of (clear-sky) Ts observations, particularly where those observations are spatially aggregated before assimilation.

This work is a first step toward evaluating LSM outputs using the remotely sensed Ts products over global land areas, and will provide useful guidance for future studies. Our comparison between the CLM4.0 modeled and MODIS observed Ts established the monthly mean differences between them, which helped to identify some deficiencies in the CLM4.0 model.

Acknowledgments:

The work of AW was supported by the Department of Science and Technology of China under Grants 2010CB428403 and the National Science Foundation of China under Grant 41275110, the work of XZ was supported by the National Science Foundation (AGS-0944101) and NASA (NNX09A021G), while the work of MB was supported by NOAA (NA13NES4400003), and the work of CSD was supported by the NASA Modeling, Analysis, and Prediction Program, and the National Climate Assessment.

References:

Angelis, C. F., G. R. McGregor, C. Kidd (2004), Diurnal cycle of rainfall over the Brazilian Amazon, *Clim. Res*., 26, 139-149.

Augustine, J. A., J. J. DeLuisi, and C. N. Long (2000), SURFRAD—A national surface radiation budget network for atmospheric research, *Bull. Am. Meteorol. Soc*., **81**, 2341-2357, doi:10.1175/1520-0477.

Baldocchi, D. et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem0-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Amer. Meteor. Soc.*, 82, 2415-2434.

Bosilovich, M., J. Radakovich , A. da Silva, R. Todling, and F.Verter (2007), Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system, *J. Meteorol. Soc. Jap*.,85A, 205-228.

Chen, Y., K. Yang, D. Zhou, J. Qin, and X. Guo (2010), Improving the Noah land surface model in arid regions with an appropriate parameterization of the thermal roughness length, *J. Hydrometeorol*., 11, 995-1006.

Garratt, J. R. (1995), Observed screen (air) and GCM surface/screen temperatures: Implications for outgoing longwave fluxes at the surface, *J. Clim.*, 8, 1360-1368.

Ghent, D., J. Kaduk, J. Remedios, J. Ardö, and H. Balzter (2010), Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter, *J. Geophys. Res*., 115, D19112, doi:10.1029/2010JD014392

Gu, L., et al. (2005), Objective threshold determination for nighttime eddy flux filtering, *Agric. For. Meteor.*, 128, 179-197.

Guo, Z., P. A. Dirmeyer, Z.-Z. Hu, X. Gao, and M. Zhao (2006), Evaluation of the second global soil wetness project soil moisture simulations: 2. Sensitivity to external meteorological forcing, *J. Geophys. Res.,* 111, D22S03, doi:10.1029/2006JD007845

Lawrence, D., et al. (2011), Parameterization improvements and functional and structural advances in version 4 of the Community Land Model, *J. Adv. Model. Earth Syst.*, 3, doi:10.1029/2011MS000045.

Lawrence, D., et al. (2012), The CCSM4 land simulation, 1850–2005: assessment of surface climate and new Capabilities, *J. Clim.*, 25, 2240-2260.

Jin, M. and S. Liang (2006), An improved land surface emissivity parameter for Land Surface Models using global remote sensing observations, *J. Clim.,* 19, 2867-2881.

Oleson, K., et al. (2010), Technical description of version 4.0 of the Community Land Model, NCAR Tech. Note NCAR/TN-4781STR, 257 pp.

Prigent, C., F. Aires, and W. B. Rossow (2003), Land surface skin temperatures from a combined analysis of microwave and infrared satellite observations for an all-weather evaluation of the differences between air and skin temperatures, *J. Geophys. Res.*, 108(D10), 4310, doi:10.1029/2002JD002301.

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2002: Part I: Forcing data and evaluations, *J. Hydrometeorol.*,7, 953–975.

Reichle, R. H., S. V. Kumar, S. P. P. Mahanama, R. D. Koster, and Q. Liu (2010),

Assimilation of satellite-derived skin temperature observations into land surface models, *J. Hydrometeorol., 11*, 1103-1122.

Salomonson, V. V., W. L. Barnes, P. W. Maymon, H. Montgomery, and H. Ostrow (1989), MODIS: Advanced facility instrument for studies of the Earth as a system, *IEEE Trans. Geosci. Remote Sens*., 27, 145-153, doi:10.1109/36.20292.

Scarino, B., P. Minnis, R. Palikonda, R. H. Reichle, D. Morstad, C. Yost, B. Shan, and Q. Liu (2013), Retrieving clear-sky surface skin temperature for numerical weather prediction applications from geostationary satellite data, *Remote Sens.,* 5, 342-366, doi:10.3390/rs5010342.

Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *J. Clim.*, **19**, 3088-3111.

Snyder, W. C., Z. Wan, Y. Zhang, and Y.-Z. Feng (1998), Classification‐based emissivity for land surface temperature measurement from space, *Int. J. Remote Sens*., 19, 2753–2774, doi:10.1080/014311698214497.

Van De Griend, A. A., and M. Owe (1993), On the relationship between thermal emissivity and the normalized difference vegetation index for nature surfaces, *Int. J. Remote Sens.,* 14**,** 1119–1131.

Wan, Z., Y. Zhang, Q. Zhang, and Z.-L. Li (2002), Validation of the land surface temperature products retrieved from Terra Moderate Resolution Imaging pectroradiometer data, *Remote Sens. Environ*., 83, 163– 180, doi:10.1016/S0034-4257(02)00093-7.

Wan, Z., Y. Zhang, Q. Zhang, and Z.-L. Li (2004), Quality assessment and validation of the MODIS global land surface temperature, *Int. J. Remote Sens.*, 25(1), 261–274, doi:10.1080/0143116031000116417.

Wan, Z., and Z.-L. Li (2008), Radiance-based validation of the V5 MODIS land-surface temperature product, *Int. J. Remote Sens*., 29, 5373–5395, doi:10.1080/01431160802036565.

Wang A. and X. Zeng (2009), Improving the treatment of vertical snow burial fraction over short vegetation in the NCAR CLM3, *Adv. Atmo. Sci.*, 26, 877-886.

Wang, A., and X. Zeng (2011), Sensitivities of terrestrial water cycle simulations to the variations of precipitation and air temperature in China, *J. Geophys. Res*., 116, D02107, doi:10.1029/2010JD014659.

Wang, K. and S. Liang (2009), Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observation at SURFRAD sites, *Remote Sens. Environ.*, 113, 1556-1565.

Xu, T., S. Liu, S. Liang, and J. Qin (2011), Improving predictions of water and heat fluxes by assimilating MODIS, *J. Hydrometeorol.*, 12, 227-244.

Yang, K., T. Koike, H. Fujii, K. Tamagawa, and N. Hirose (2002), Improvement of surface flux parameterizations with a turbulence‐related length, *Q. J. R. Meteorol. Soc.*, 128, 2073‐2087, doi:10.1256/003590002320603548.

Yang,K., et al. (2008), Turbulent flux transfer over bare soil surfaces: Characteristics and parameterization, *J. Appl. Meteorol. Clim.*, **40**, 276-290.

Zeng X., and A. Wang (2007), Consistent parameterization of above-canopy

turbulence for sparse and dense canopies in land models, *J. Hydrometeorol.*, 8, 730-737.

Zeng, X. and R. E. Dickinson (1998), Effect of surface sublayer on surface skin temperature and fluxes, *J. Clim.*, 11, 537-550.

Zeng, X., Z. Wang, and A. Wang (2012), Surface skin temperature and the interplay between sensible and ground heat fluxes over arid regions, *J. Hydrometeorol*. 13, 1359-1370.

Zheng, W., H. Wei, Z. Wang, X. Zeng, J. Meng, M. Ek, K. Mitchell, and J. Derber (2012), Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation, *J. Geophys. Res.*, 117, D06117, doi:10.1029/2011JD015901.

Table captions

Table 1. Information of four stations used in this study.

Table 2. Monthly mean Ts differences between MODIS, CLM-C and CLM-N versus in situ observations over four stations at four satellite overpass times in July, 2003. Only the values under clear-sky conditions as indicated by the MODIS Ts data are used. The corresponding biases between Tair and downward shortwave radiation (SWdn) between CLM forcing and in-situ measurements are also shown in the last two columns. Biases that are statistically significant at the 1% level based on the Student's t-test are indicated in bold.

Table 3. Ratios of the standard deviations (STD) of Ts differences (STDd) between model or MODIS results and in situ observations to the STD of in-situ observations (STDo) over four stations at four satellite overpass times in July 2003.

Table 4. Monthly Ts differences (K) averaged over Northern Hemisphere (NH) and Southern Hemisphere (SH) land grid cells between CLM-C and MODIS, and between CLM-N and MODIS in January and July 2003, respectively. At each MODIS satellite overpass time, only the grid cells meeting two criteria are used to compute monthly Ts in CLM: a) bare fraction (BF) is greater than 30%; and b) MODIS clear-sky fraction (CF) is greater than 50% for at least 10 days in the month.

Figure captions

Figure 1. CLM4.0 grid box number percentages over land (excluding the Antarctic) versus clear-sky percentages using results from each overpass for each day for the whole month in January and July, 2003.

Figure 2. Monthly Ts differences between CLM-C and MODIS at four overpass times in July 2003. At each overpass time, CLM-C monthly Ts values are computed only for grid boxes with MODIS clear-sky fraction > 50% for at least 10 days in the month. The areal weighted values over each hemispheric land areas are also shown in the figure.

Figure 3. Hemisphere mean Ts differences between CLM-N and CLM-C versus bare soil fraction in 5% intervals at four satellite overpass times averaged in January and July 2003. NH and SH denote Northern and Southern Hemispheres, respectively.

Figure 4. Global distribution of Ts differences between CLM-N and CLM-C at a) 10:30am; b) 1:30pm, and between CLM-C and MODIS at c) 10:30am; d) 1:30pm in July 2003. At each satellite overpass time, monthly Ts is computed over grid boxes with bare soil fraction greater than 30% and MODIS clear-sky fraction greater than 50% for at least 10 days in the month.

Figure 5. Similar as Figure 4 but for January 2003.

Figure 6. a) Monthly averaged Ts at two grid boxes at the four satellite overpass times and b) the Ts differences between these two boxes (centered around 29°N/23°E, and 29°N/10°W) from CLM-C and MODIS. Both boxes are located in the Sahara Desert with bare soil fraction greater than 90%.

Table 1. Information of four stations used in this study.

Station name	Location		Surface emissivity	Data sources	References
	Lat(°N)	Lon(°E)			
Desert Rock	36.62	-116.02	0.96	SURFRAD	*Augustine et al.* 2000
Colorado	40.13	-105.24	0.98	SURFRAD	*Augustine et al.* 2000
Tongyu	44.41	122.87	0.96	CEOP	*Yang et al*. 2008
Gaize	32.3	84.5	0.91	CEOP	*Chen et al*. 2010

Table 2. Monthly mean Ts differences between MODIS, CLM-C and CLM-N versus in situ observations over four stations at four satellite overpass times in July 2003. Only the values under clear-sky conditions as indicated by the MODIS Ts data are used. The corresponding biases between Tair and downward shortwave radiation (SWdn) between CLM forcing and in-situ measurements are also shown in the last two columns. Biases that are statistically significant at the 1% level based on the Student's t-test are indicated in bold.

		Ts diff. (K)			Tair diff. (K)	SWdn diff. (W/m^2)
		MODIS	CLM-C	CLM-N		
Desert Rock	1:30a	**-4.14**	**-6.47**	**-5.69**	**-10.31**	0.
	10:30a	**2.23**	**-3.79**	-1.85	**-3.02**	**-142.46**
	1:30p	**-1.30**	**-4.35**	-1.61	**-1.75**	**-154.26**
	10:30p	**-4.17**	**-5.72**	**-4.92**	**-8.47**	0
Colorado	1:30a	**-4.07**	**-5.22**	**-4.78**	**-9.98**	0
	10:30a	**2.27**	**-7.02**	**-6.83**	**-4.34**	**-207.06**
	1:30p	**-1.26**	**-5.95**	**-5.53**	**-3.65**	-77.63
	10:30p	**-4.26**	**-5.07**	**-4.55**	**-7.99**	0
Tongyu	1:30a	-2.55	-0.36	-0.15	-0.87	0
	10:30a	-2.30	**-5.31**	-4.86	**-4.15**	**-215.74**
	1:30p	-1.15	-2.43	-2.03	-1.94	-78.54
	10:30p	**-1.93**	0.19	0.40	0.71	0
Gaize	1:30a	**-3.51**	**-2.27**	-1.23	**-3.76**	0
	10:30a	**10.61**	**-8.76**	**-7.06**	**-9.25**	**-215.91**
	1:30p	**1.92**	**-11.41**	**-8.91**	**-7.79**	**-186.43**
	10:30p	**-5.21**	**-2.83**	-1.59	**-3.99**	0

Table 3. Ratios of the standard deviations (STD) of Ts differences (STDd) between model or MODIS results and in situ observations to the STD of in-situ observations (STDo) over four stations at four satellite overpass times in July 2003.

		STDd/STDo		
		MODIS	CLM-C	CLM-N
Desert Rock	1:30am	0.50	0.53	0.55
	10:30am	1.05	0.79	0.85
	1:30pm	0.96	1.18	1.33
	10:30pm	1.56	0.56	0.58
Colorado	1:30am	0.82	0.91	0.91
	10:30am	1.06	0.95	0.95
	1:30pm	0.70	0.97	0.98
	10:30pm	1.05	0.83	0.83
Tongyu	1:30am	0.62	0.20	0.20
	10:30am	1.21	1.02	1.02
	1:30pm	1.04	1.05	1.05
	10:30pm	0.73	0.51	0.52
Gaize	1:30am	1.50	1.02	1.00
	10:30am	0.91	0.81	0.79
	1:30pm	1.02	0.74	0.71
	10:30pm	1.81	0.95	0.94

Table 4 Monthly Ts differences (K) averaged over Northern Hemisphere (NH) and Southern Hemisphere (SH) land grid boxes between CLM-C and MODIS, and between CLM-N and MODIS in January and July 2003, respectively. At each MODIS satellite overpass time, only the grid boxes meeting two criteria are used to compute monthly Ts in CLM: a) bare fraction (BF) is greater than 30%; and b) MODIS clear-sky fraction (CF) is greater than 50% for at least 10 days in the month.

		SH		NH	
		CLM-C& MOD	CLM-N& MOD	CLM-C& MOD	CLM-N& MOD
January	10:30am	-7.73	-6.31	-6.50	-6.14
	1:30pm	-4.36	-1.98	-2.65	-1.76
July	10:30am	-5.65	-5.27	-5.60	-4.47
	1:30pm	-3.69	-2.86	-0.75	1.25

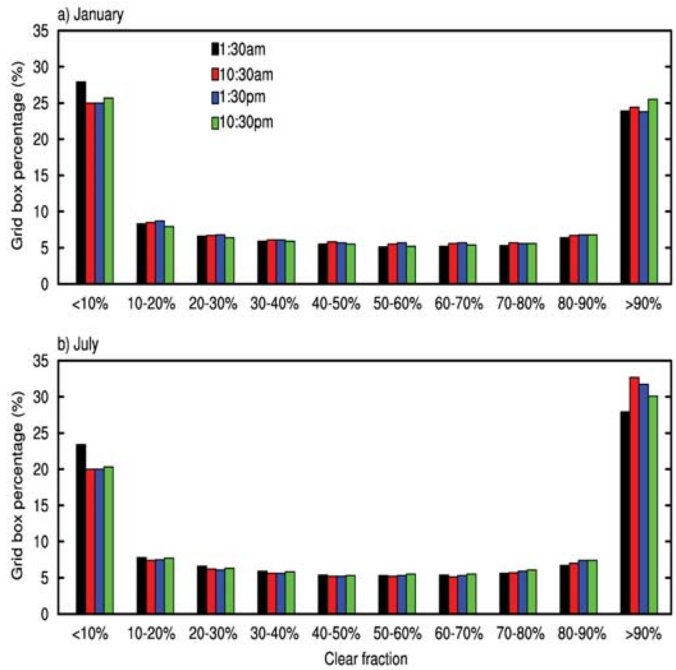

Figure 1. Distribution of global land CLM4.0 grid boxes (excluding the Antarctic) by clear-sky fraction, using results from each overpass for each day for the whole month in January and July, 2003.

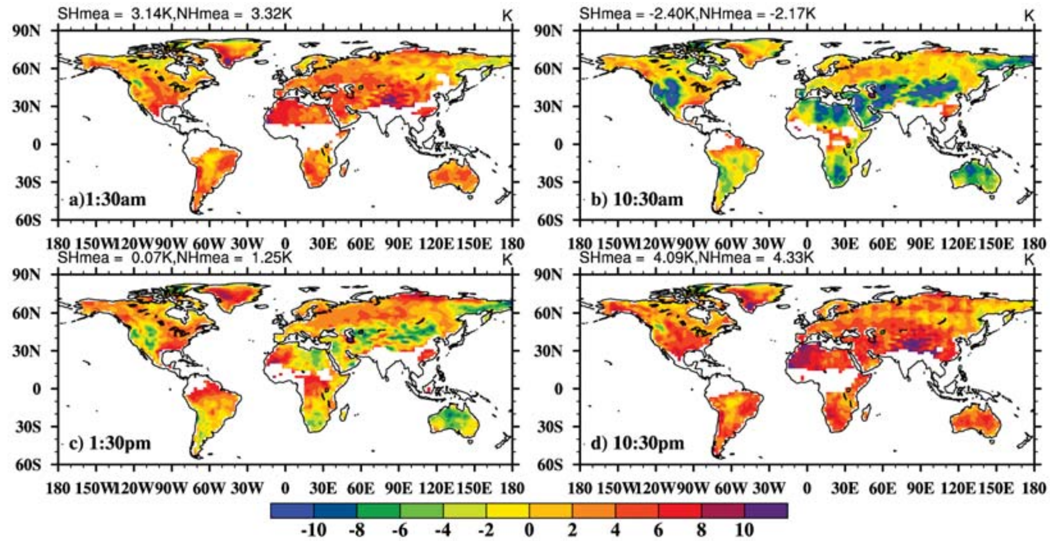

Figure 2. Monthly Ts differences between CLM-C and MODIS at four overpass times in July 2003. At each overpass time, CLM-C monthly Ts values are computed only for grid boxes with MODIS clear-sky fraction > 50% for at least 10 days in the month. The areal weighted values over each hemispheric land areas are also shown in the figure.

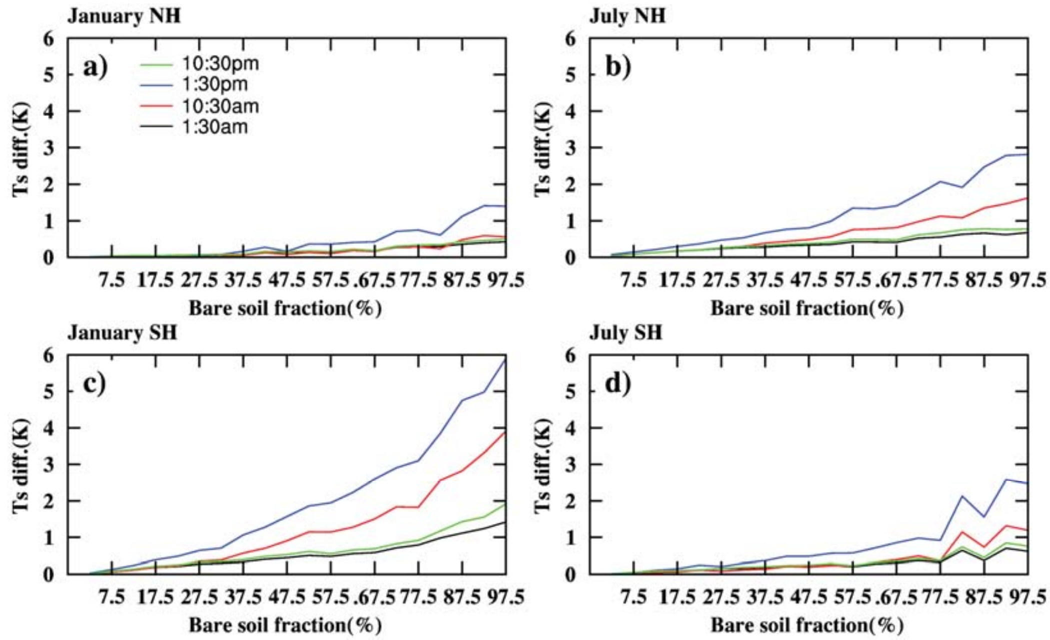

Figure 3. Hemisphere mean Ts differences between CLM-N and CLM-C versus bare soil fraction in 5% intervals at four satellite overpass times averaged in January and July 2003. NH and SH denote Northern and Southern Hemispheres, respectively.

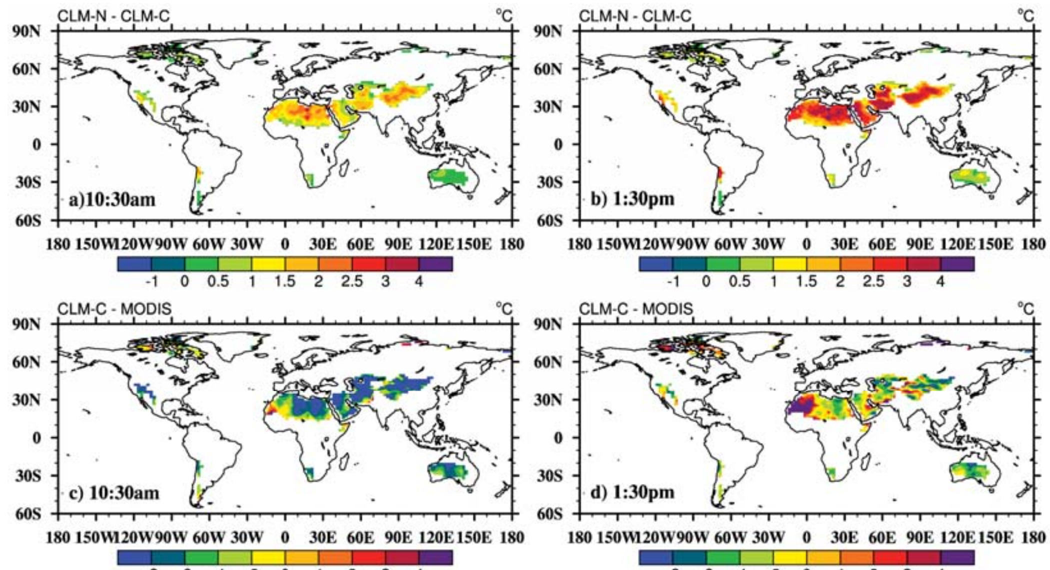

Figure 4. Global distribution of Ts differences between CLM-N and CLM-C at a) 10:30am; b) 1:30pm, and between CLM-C and MODIS at c) 10:30am; d) 1:30pm in July 2003. At each satellite overpass time, monthly Ts is computed over grid boxes with bare soil fraction greater than 30% and MODIS clear-sky fraction greater than 50% for at least 10 days in the month.

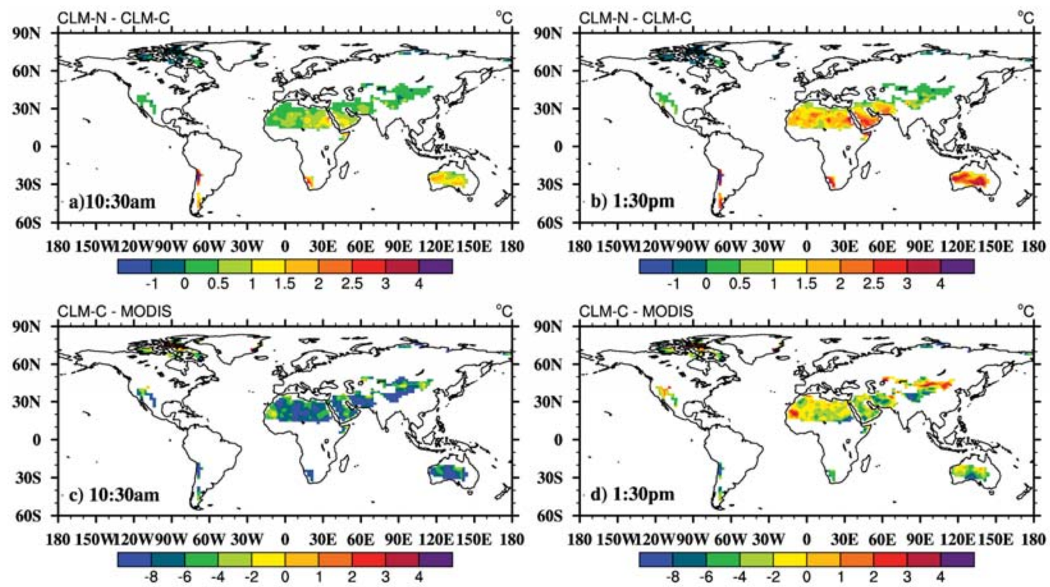

Figure 5. Similar as Figure 4 but for January 2003.

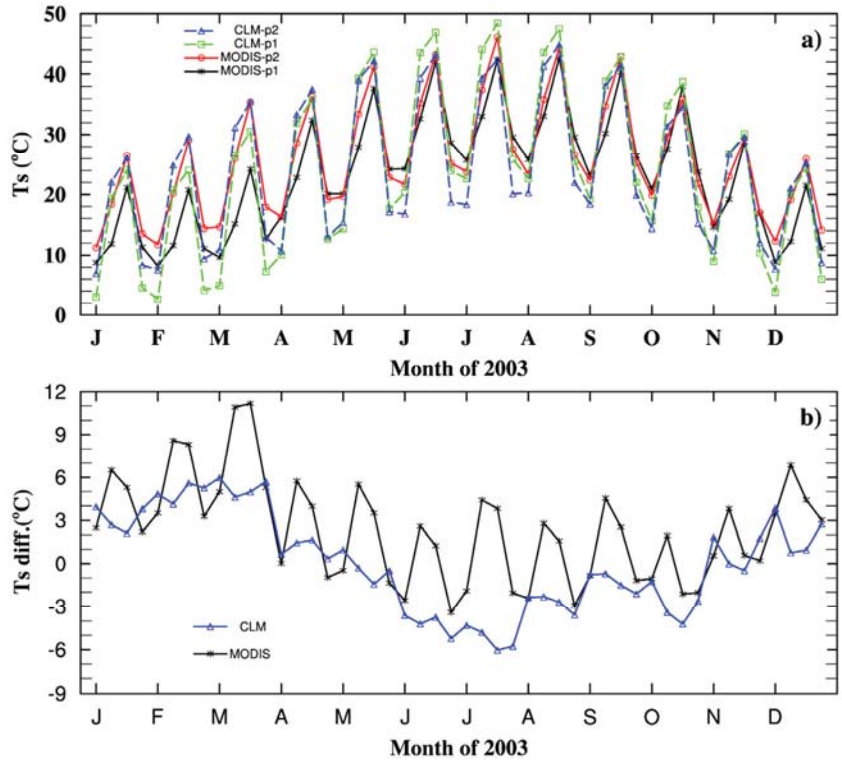

Figure 6. a) Monthly averaged Ts at two grid boxes at the four satellite overpass times and b) the Ts differences between these two boxes (centered around 29°N/23°E and 29°N/10°W) from CLM-C and MODIS. Both boxes are located in the Sahara Desert with bare soil fraction greater than 90%.